

Citation for published version:

Kjeldsen, TR 2015, 'How reliable are design flood estimates in the UK?', *Journal of Flood Risk Management*, vol. 8, no. 3, pp. 237-246. <https://doi.org/10.1111/jfr3.12090>

DOI:

[10.1111/jfr3.12090](https://doi.org/10.1111/jfr3.12090)

Publication date:

2015

Document Version

Peer reviewed version

[Link to publication](#)

This is the peer reviewed version of the following article: Kjeldsen, TR 2015, 'How reliable are design flood estimates in the UK?' *Journal of Flood Risk Management*, vol 8, no. 3, pp. 237-246., which has been published in final form at 10.1111/jfr3.12090. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

How reliable are design flood estimates in the UK?

Thomas Rodding Kjeldsen

Department of Architecture and Civil Engineering, University of Bath, Bath, BA2 7AY, UK.
t.r.kjeldsen@bath.ac.uk

Abstract Design flood estimates in the UK are routinely obtained by using the improved Flood Estimation Handbook (FEH) statistical procedure. This paper presents a practical framework for assessing the uncertainty associated with estimates of the index flood (QMED) obtained for a range of commonly encountered problems: the gauged, the ungauged and the sparsely gauged catchment. An assessment is presented of the uncertainty of design flood estimates when estimated at ungauged catchments for a range of return periods. The results show that the inclusion of data from nearby gauged catchments increases the reliability of the estimates when compared to an automated application of the improved FEH methods relying on catchment descriptors only.

1. INTRODUCTION

Design flood estimates are required for a range of engineering and planning purposes. Most commonly these estimates are obtained from statistical models fitted to annual maximum (AMAX) series of peak flow for either single site or a standardised (regional) combination of sites depending on the data availability at the site of interest (Stedinger et al., 1993). In the United Kingdom the first comprehensive country-wide methodology for flood frequency estimation was the Flood Studies Report (FSR) published by the Natural Environment Research Council (NERC, 1975). The FSR method was based on the index flood model previously developed in the USA (Dalrymple, 1960) and divided the British Isles into 10 geographical regions and one for Ireland. For each region, a catchment specific flood frequency curve, here understood as the probabilistic model linking flood magnitude (m^3/s) to exceedance probability (the inverse of return period), could be derived as the product of a regional dimensionless growth factor and an estimate of the index flood (the mean annual maximum flood) obtained either directly from observations or via a regression model based

on catchment characteristics (catchment area, soil type, annual average rainfall etc.). Following methodological advances in regional flood frequency estimation, and in particular the introduction of the method of L moments (Hosking and Wallis, 1993) and the region of influence approach (Burn, 1990), the Institute of Hydrology (1999) published revised procedures in the Flood Estimation Handbook (FEH). The FEH methodology introduced a number of new features into practical flood frequency estimation, including the use of: i) the concept of hydrological similarity for use in the formation of catchment specific growth curves through pooling groups, rather than the regional values presented in the FSR, ii) the median annual maximum flood as a measure of the index flood, as the median was considered less sensitive than the mean annual flood to outliers in relatively short data series, iii) formalised method of data transfer from nearby hydrologically similar (donor) sites when estimating the index flood in ungauged catchments, and iv) catchment descriptors derived automatically from gridded electronic datasets rather than catchment characteristics extracted manually from 1:50,000 scale paper maps. The index flood model and procedures similar to the FEH have also been developed for other parts of the world, including: Africa (Mkhandi et al., 2000), Asia (Yang et al., 2010), and parts of Europe (Castellarin et al., 2012). More recently, an improved set of FEH procedures for use in the United Kingdom were developed and published by the Environment Agency (2008) and implemented in the WINFAP-FEH v3 software (WHS, 2009). The improved method has retained the index flood model and the use of the hydrological similarity method, but important improvements were made to the underlying statistical models for estimating both the index flood (including donor transfer) and the dimensionless growth curve.

Depending on data availability, the improved FEH methodology provides a range of methods that can be used for calculating design floods, but little advice is provided on how to assess the uncertainty of these estimates; e.g. the 95% or 68% confidence intervals are commonly used as they are easily obtained as plus-minus one and two times the standard deviation. This is particularly true for the case where data transfer from nearby gauged donor sites is used in an attempt to get more precise estimates of the index flood (defined as the median annual maximum peak flow) at ungauged catchments. In the case where no local data are available, *the index flood* is obtained directly from catchment descriptors using an established regression model.

Recent years have seen a growing call for risk and uncertainty to be given a more central role in flood management (e.g. Reed, 2002; Pappenberger and Beven, 2006; Hall, 2011). This gap

between the requirements of managers and the information and guidance that can currently be provided by engineers and hydrologists needs to be closed. The results presented in this paper provide a new set of methods and results that enables the uncertainty in design flood estimates from the recently published Improved FEH statistical method (Environment Agency, 2008) to be assessed. The paper will primarily discuss the accuracy of the index flood variable when estimates are made at sites where no or only few flood flow data are available (i.e. ungauged sites). The paper will also provide an assessment of the expected accuracy of the final design flood estimate at ungauged sites, across a range of return periods, when the improved FEH methods are used.

2. THE IMPROVED FEH STATISTICAL METHOD

The improved FEH procedure (Environment Agency, 2008) is primarily based on the analysis of annual maximum (AMAX) series of instantaneous peak flow, and while peaks-over-threshold (POT) series might be used in certain circumstances, the discussions in this paper are confined to AMAX data only. The discussion of uncertainty of design flood estimates in the following is predicated on a short introduction of the FEH methodology, but for a more detailed description of current procedures for flood estimation the reader is advised to consult the Environment Agency R&D report (Environment Agency, 2008).

Consider the formula for the T -year design event flow Q_T when estimated using a Generalised Logistic (GLO) distribution

$$Q_T = \mu + \frac{\alpha}{\kappa} \left(1 - (T - 1)^{-\kappa} \right) \quad (1)$$

where μ , α , and κ are the location, scale and shape parameters of the GLO distribution and T is the return period, measured in years. The estimator can be re-written in the format of the index flood method as

$$Q_T = \mu \left[1 + \frac{\beta}{\kappa} \left(1 - (T - 1)^{-\kappa} \right) \right] = \mu \cdot z_T \quad (2)$$

where $\beta = \alpha/\mu$. The dimensionless growth factor is denoted z_T and μ is the location parameter and also known as the index flood. In the FEH the index flood is defined as the median annual maximum flood, which represents the design flood with a return period of 2

years. The location parameter, μ , is estimated as the sample median and is therefore traditionally denoted *QMED* in FEH terminology. The scale and shape parameters are estimated using the method of L moments as described by the Environment Agency (2008). In particular, L skewness defines the shape parameter κ as

$$\kappa = -(\text{L skewness}). \quad (3)$$

When conducting a flood frequency analysis at a gauged site, the three model parameters (μ , β and κ) can be estimated directly from the available AMAX data. When a design flood estimate is required at an ungauged site, the growth curve, z_T , is estimated by transferring data from other gauged catchments that are considered hydrologically similar to the site of interest. In practice this is done through the formation of a pooling group, normally consisting of the, typically 15-17, most similar catchments. For an ungauged site the index flood is estimated by using a standard regression model linking the index flood to a set of catchment descriptors. Details of how to form pooling group and estimate model parameters using the method L-moments is provided by the Environment Agency (2008) and Kjeldsen and Jones (2009).

3. MEASURING UNCERTAINTY IN DESIGN FLOODS

Before discussing the uncertainty of the FEH design flood estimates, it is convenient to discuss how uncertainty is measured. It is common to use the variance or the square root of the variance, i.e. the standard deviation, to quantify the uncertainty of a random variable such as design flood estimates. When a random variable, say x , is assumed normally distributed, the standard deviation, s_x , is often used to provide the 68% or 95% confidence intervals as

$$x \pm s_x \quad \text{68\% confidence interval}$$

$$x \pm 2s_x \quad \text{95\% confidence interval}$$

Often in hydrology, it is not the random variable itself, but rather the log-transformed variable that is assumed normally distributed. In such cases, the random variable $x = \ln q$, can be transformed using the exponential function, $q = \exp(x)$. By applying the exponential function to the confidence intervals limits, these can be defined as

$$\left[e^x e^{-s_x}; e^x e^{s_x} \right] = \left[\frac{q}{fse}; q \cdot fse \right] \quad 68\% \text{ confidence interval}$$

$$\left[e^x e^{-2s_x}; e^x e^{2s_x} \right] = \left[\frac{q}{fse^2}; q \cdot fse^2 \right] \quad 95\% \text{ confidence interval}$$

On the right-hand side of the equations the factorial standard error, fse , has been defined by applying the exponential function to the estimate of the standard deviation of the log-transformed variable, i.e.

$$fse = e^{s_x} \quad (4)$$

This is relevant, especially when considering the uncertainty of $QMED$ for estimates obtained using the regression model linking $QMED$ to catchment descriptors, as this equation was developed by considering the model residuals to be normally distributed when modelling the log-transformed $QMED$ values. Therefore, in the following analysis the uncertainty for all considered cases will be discussed and quantified in terms of the fse .

The uncertainty of the design flood estimates obtained by Eq. (2) can be estimated using a first order approximation of the variance of a product as

$$\text{var}\{\hat{Q}_T\} = z_T^2 \text{var}\{\hat{\mu}\} + \mu^2 \text{var}\{\hat{z}_T\} + 2\mu z_T \text{cov}\{\hat{\mu}, \hat{z}_T\} \quad (5)$$

which shows that the total uncertainty consists of a contribution from the uncertainty of the index flood, the growth curve and the covariance between the two. In a series of papers by Kjeldsen and Jones (2004; 2006) methods for estimating the uncertainty of single site and pooled estimates of design floods were presented based on Eq. (5) assuming a GLO distribution. While in some specific cases relatively simple analytical expressions of uncertainty in the index flood can be derived, the contribution from the growth curve and the covariance terms are much more complicated. In particular, the uncertainty of the growth curve needs to consider the effects of dependence between AMAX series across sites included in the pooling groups. To the author's knowledge, these methods have so far enjoyed limited practical use.

In the following, simple and practical expressions of uncertainty of the index flood ($QMED$) will be derived, followed by a more general and empirical assessment of the uncertainty associated with design flood estimates of a higher return period at ungauged sites. The latter

estimates will provide a more directly and easily applicable set of guidelines for levels of uncertainty than those provided by Eq. (5). For a more detailed discussion of methods for assessing the uncertainty of design events obtained from the FEH at gauged sites using single site and pooled analysis see Kjeldsen and Jones (2004, 2006) and Kjeldsen *et al.* (2013) for a broader discussion of uncertainty in flood frequency estimation.

4. ESTIMATION OF THE INDEX FLOOD (QMED)

The index flood, or *QMED*, plays an important part of the FEH statistical method. In the following section, useful expressions of the factorial standard error will be derived for a range of situations which will commonly occur when applying the FEH methodology in practise.

The gauged site

First, consider the case where a sample of AMAX data is available at the site of interest (single site analysis), in which case *QMED* is the middle-ranking value in an ordered sample with n observations and defined as:

$$QMED = \begin{cases} q_{[m]} & \text{where } m = (n+1)/2 \text{ for } n \text{ odd} \\ (q_{[m]} + q_{[m+1]})/2 & \text{where } m = n/2 \text{ for } n \text{ even} \end{cases} \quad (6)$$

It can be shown (e.g. Cox and Hinkley, 1974) that the variance of the median (here denoted $var\{QMED\}$ or s_Q^2) can be approximated for large sample sizes as

$$s_Q^2 \approx \frac{1}{4nf^2(F^{-1}(0.5))} \quad (7)$$

where n is the number of AMAX events in the sample, f is the probability density function (pdf) of the AMAX series, and $F^{-1}(0.5)$ is the median quantile of the distribution, i.e. *QMED*. The discussion here assumes that the AMAX events follow a Generalised Logistic (GLO) distribution. Other distributions such as the Generalised Extreme Value (GEV) or Gumbel (EV1) distributions can be used, leading to results which will be different from the once presented here. Note that the expression in Eq. (7) is derived for use with large sample sizes, and thus might provide unreliable estimates for small samples.

In order to use the fse to quantify the uncertainty it is necessary to first consider the standard deviation of the log-transformed median, and not just the median as in Eq. (7). As described by Kjeldsen and Jones (2006), using a first order approximation the standard deviation of the log-transformed median can be derived from Eq. (7) as

$$s_x = \frac{2\beta}{\sqrt{n}} \quad (8)$$

where β is the scale parameter of the GLO distribution, estimated using for example the L-moment ratios (equation 15.8 in FEH Vol.3). Combining the expression in Eq. (8) with Eq. (4), the fse of QMED when derived from at-site AMAX data is defined as

$$fse = e^{(2\beta/\sqrt{n})} \quad (9)$$

The fse depends on the length of the AMAX record, n , showing that the more AMAX data are available, the more confident the analyst can be in the estimated value of QMED, which is reassuring.

Notice that a suitable transformation of the standard deviation in Eq. (8) will result in a constant value

$$\frac{\sqrt{n}s_x}{\beta} = 2 \quad (10)$$

However, the quality of first order approximations used for deriving Eq. (8) are known to depend on the degree of non-linearity of the quantile function (Eq. 2) and the sample size (e.g. Kjeldsen and Jones, 2004). An assessment of the performance of Eq. (8) for a range of sample sizes ($n = 5, 10, 15, 25, 50, 100$) and skewness values ($-0.5 < \kappa < 0.5$, by 0.05) was carried out through a Monte Carlo simulation study. For each combination of sample size and skewness, 10000 random samples were generated from a GLO distribution, and the standard deviation of the log-transformed median estimated. For each sample size, a 3rd order polynomial was fitted to the generated data,

$$\frac{\sqrt{n}s_x}{\beta} = \sum_{i=0}^3 a_{(i)} \kappa^i \quad -0.5 < \kappa < 0.5 \quad (11)$$

expressing the transformed standard deviation as a function of the shape parameter. A comparison of the MC generated results and the fitted 3rd order polynomial are shown in Figure 1, and the coefficients $a_{(i)}$, of the polynomial listed in Table 1

FIGURE 1

TABLE 1

Figure 1 shows that for negative values of the κ -parameter (positive L skewness) the approximations perform reasonably well. However, for negatively skewed distributions and for small sample sizes, the approximations deviate considerably from the constant value of 2. Thus, it is not advisable to estimate the *fse* of *QMED* for short samples, say less than 20 AMAX events, using the Eqs. (8) and (9). The polynomial fitted to the Monte Carlo results will be used later to evaluate the uncertainty of the index flood for the case where a short series of AMAX events are combined with a generalised estimate of *QMED* based on the use of FEH catchment descriptors.

The ungauged site

When no AMAX data are available at the site of interest, hydrologists have to employ indirect methods to obtain useful estimates of *QMED*. Here uncertainty measures are presented for two cases: i) where *QMED* is estimated from catchment descriptors only, and ii) the estimate of *QMED* from catchment descriptors is adjusted using data-transfer from a nearby gauged donor site.

Using catchment descriptors only

The improved FEH statistical method (Environment Agency, 2008) presented a linear regression model linking the log-transformed value of *QMED* to a set of four catchment descriptors as

$$\ln QMED = 2.1170 + 0.8510 \ln[AREA] + 1.8734 \left[\frac{1000}{SAAR} \right] + 3.4451 \ln[FARL] - 3.080 BFIHOST^2 \quad (12)$$

where *AREA* is the catchment area (km²), *SAAR* is the standard annual average rainfall as defined for the period 1961-90 (mm), *FARL* is an index of flood attenuation from upstream lakes and reservoirs, and *BFIHOST* is the baseflow index as defined from Hydrology of Soil Types (HOST) data (Boorman et al., 1995). All four descriptors are available on the FEH CD-ROM v3 (CEH, 2009).

When using a regression model, like Eq. (12), to predict *QMED* in an ungauged catchment, the total prediction variance is a rather complicated expression involving matrixes and access to all the data used in the development of the model itself. However, given the relatively large number of gauged catchments (602) used in the development of Eq. (12), the prediction variance is dominated by the variance of the regression model residuals, which was reported by the Environment Agency (2008) as $s^2 = 0.1286$. Thus, the *fse* for this model is easily defined as

$$fse \approx e^s = e^{\sqrt{0.1286}} = 1.431 \quad (13)$$

Note that unlike the *fse* for the gauged case in Eq. (9), when using catchment descriptors only to estimate *QMED*, *fse* is considered identical for all catchments.

Using data transfer from a donor catchment

The FEH strongly recommends that estimates of *QMED* using catchment descriptors should be used only as a last resort; a view that was reinforced by Reed (2002). If no at-site data are available, hydrologists are advised to use data transfer from suitable donor sites to improve the estimate of *QMED* obtained from catchment descriptors. An important feature of the improved FEH statistical method is the introduction of a revised procedure to be used for data transfer. The rationale behind the revised procedure is that local data are used to compensate for the inability of the model in Eq. (12) to estimate *QMED* resulting from the use of simplistic and lumped catchment descriptors to represent complex catchment hydrology. Research leading to the improved procedure (Kjeldsen and Jones, 2007) found that geographical proximity was more important than assumed hitherto, and should be used to determine the weight assigned to donor sites. Consequently, a revised data transfer procedure was introduced where the adjustment factor is down-scaled as the geographical distance

increases between the centroids of the catchments draining to the site of interest and the donor site, respectively. Analytically, the revised scheme is defined as

$$QMED_{s,adj} = QMED_{s,cds} \left(\frac{QMED_{g,obs}}{QMED_{g,cds}} \right)^\alpha \quad (14)$$

where the subscripts are defined as follows; s and g : the ungauged subject site and the gauged sites, respectively; cds : catchment descriptor estimates at the gauged and ungauged sites; obs : the observed value at the gauged site; adj : the adjusted value at the subject site. The weighting factor α is a function of geographical distance and defined as

$$\alpha = 0.4598e^{(-0.0200d_{sg})} + (1 - 0.4598)e^{(-0.4785d_{sg})} \quad (15)$$

Where d_{gs} is the geographical distance (in km) between the centroids of the subject and donor catchments. To define the fse of $QMED$ estimated using Eq. (14), it is necessary first to consider the variance of the log-transformed $QMED$ which was provided by Kjeldsen and Jones (2007) as

$$\text{var}\{\ln QMED_{s,adj}\} = s^2(1 - \alpha^2) + \alpha^2 s_{x,donor}^2 \quad (16)$$

where $s_{x,donor}^2$ is the sampling variance of the log-transformed $QMED$ at the donor site as calculated using Eq. (8), s^2 is the residual variance for the $QMED$ equation (Eq.13), and α is defined in Eq. (15). The total uncertainty of $QMED$ consists of contributions from the estimates obtained using catchment descriptors at the subject and donor sites ($QMED_{s,cds}$ and $QMED_{g,cds}$) and from the uncertainty of the estimate of $QMED_{g,obs}$ obtained from the data at the donor site. From Eq. (16) the fse for $QMED$ obtained from data transfer is defined as

$$fse = e^{\sqrt{s^2(1-\alpha^2) + \alpha^2 s_{x,donor}^2}} \quad (17)$$

If a relatively large number of AMAX data (say more than 15) are available, and the distance between the donor and subject site is more than a few kilometres, then the second term in Eq. (17), is very likely to be much smaller than the first term and fse can be reduced to

$$fse \approx e^{(s\sqrt{1-\alpha^2})} \quad d > 0 \text{ km} \quad (18)$$

which for distances between catchment centroids larger than zero will always give lower estimates of fse for the index flood than corresponding estimates obtained directly from

catchment descriptors only (Eq. 13). As mentioned above, the donor transfer method can be interpreted as using local data to compensate for the inability of the *QMED* equation in Eq. (12) and the lumped FEH catchment descriptors to adequately capture all the local factors controlling catchment flood response. A more detailed discussion of the use of data transfer is provided by Kjeldsen and Jones (2007).

The sparsely gauged site

In some cases a short sample of AMAX data might be available at the site of interest and it can then be beneficial to estimate *QMED* as a weighted average of the sample *QMED* from the observations and the more generalised estimate obtained from catchment descriptors (Eq. 12). A weighted average could take the form as

$$\ln QMED_{s,adj} = \omega \ln QMED_{s,cds} + (1 - \omega) \ln QMED_{s,obs} \quad (19)$$

where the subscripts have the same meaning as before, and ω is a weight between 0 and 1. The variance of the estimates from Eq. (19) above can be expressed as

$$\begin{aligned} \text{var}\{\ln QMED_s\} &= \omega^2 s^2 + (1 - \omega)^2 s_x^2 \\ &= s_x^2 + \omega^2 s_x^2 + \omega s^2 - 2\omega s_x^2 \end{aligned} \quad (20)$$

assuming independence between the regression model estimate and the at-site estimate of *QMED*. By considering the minimal prediction variance using the analytical framework outlined by Kjeldsen and Jones (2007) and aiming to obtain the set of weights that gives the minimum prediction variance gives

$$\omega = \frac{s_x^2}{s^2 + s_x^2} \quad (21)$$

where s_x^2 is the variance of the (log-transformed) observed *QMED* and s^2 is the variance of the prediction from the regression model. Substituting Eq. (21) back into Eq. (20), the variance of the log-transformed estimate of *QMED* is then given as

$$\begin{aligned} \text{var}\{\ln QMED_s\} &= s_x^2 + \omega^2 s_x^2 + \omega^2 s^2 - 2\omega s_x^2 \\ &= \left(\frac{s^2}{s_x^2 + s^2} \right)^2 s_x^2 + \left(\frac{s_x^2}{s_x^2 + s^2} \right)^2 s^2 \end{aligned} \quad (22)$$

From Eq. (22) the corresponding fse is given as

$$fse = e^{\frac{s_x s}{\sqrt{s_x^2 + s^2}}} \quad (23)$$

As previously mentioned, the variance of the sample median can be poorly estimated using the large sample approximation in Eq. (8) where sample sizes are small and when κ is positive (negative L skewness). This might result in counterintuitive results when applied in the context of the weights presented in Eq. (21) when only a short AMAX record is available at the site of interest. Alternatively, a more believable estimate of s_x might be gained by using the 3rd order polynomial listed in Table 1, with an estimate of the GLO shape parameter, κ , obtained using a regional frequency method such as the pooled flood frequency method outlined in Kjeldsen and Jones (2009) for ungauged or enhanced single site analysis.

5. A PRACTICAL EXAMPLE OF ESTIMATING *QMED* AND ITS UNCERTAINTY

Essentially, Eqs. (9), (13), (18) and (23) provide estimates of fse for each of the four methods for estimating the index flood depending on data availability (gauged, regression model only, regression model combined with data transfer, and combining at-site and regional data). To further illustrate their use, consider a simple example where *QMED* and the associated fse values are estimated for a particular site using each of the four methods. The catchment characteristics for the subject and donor site are shown in Table 2 and the results summarised in Table 3.

A series of AMAX events is available at the flow gauging station on the River Yscir at Pont-ar-yscir, National River Flow Archive (NRFA) station no. 56013, for the period 1972-2009; a total of 36 observations as shown in Figure 2. The length of the data record available at this station represents a typical situation in applied flood hydrology in the UK.

FIGURE 2 Annual maximum series of peak flow at gauging station Yscir at Pont-ar-yscir (56013).

From the AMAX events, the median annual maximum flood ($QMED$) is estimated to be $36.6 \text{ m}^3/\text{s}$ and the associated fse is derived via Eq. (9) to be 1.08. The FEH catchment descriptors for the site are shown in Table 2, and using the regression model the $QMED$ is estimated to be $31.6 \text{ m}^3/\text{s}$ with an associated fse value of 1.43. Clearly, estimates derived from actual observations are preferable to getting estimates via catchment descriptors only.

TABLE 2: FEH catchment descriptors at subject and donor site

TABLE 3: Summary of $QMED$ calculations

Based on geographical distance between catchment centroids, the nearest suitable donor site is located on the River Honddu at The Forge (NRFA no. 56003); suitability of the donor site is here defined in terms of geographical proximity and similarity in catchment descriptor space. The catchment descriptors of the donor site are shown in Table 2. The geographical distance between the centroids of the two catchments draining the subject and donor site, respectively, is 5.0 km. At the donor site a series of AMAX events are available for the period 1963-1983 and the associated value $QMED_{g,obs}$ is estimated to be $23.5 \text{ m}^3/\text{s}$. The corresponding estimate from the regression model at the donor site is $23.9 \text{ m}^3/\text{s}$. Using Eq. (14) and Eq. (18) the adjusted estimate of $QMED$ at the subject site is $31.3 \text{ m}^3/\text{s}$ with an fse of 1.37.

These results show that estimating $QMED$ from a long record should be the preferred option where possible as reflected in the low value of fse . If the site is ungauged, the use of data transfer from a nearby catchment will increase the confidence of the estimated $QMED$ when compared to an estimate directly from use of catchment descriptors only, even if the value of the $QMED$ estimate itself does not change very much in this case.

Finally, the situation is considered where a short series of annual maximum peak flow events is available at a site of interest (the sparsely gauged site). For the purpose of this example, consider that AMAX events are available only for the last five years at the site of interest; gauging station 56013. This short series includes the largest event recorded at this station during the summer 2007 flood, which has a pronounced effect on the estimated value of the L skewness (0.77) in the short record. This value is beyond the range of values defined for the

GLO distribution. An alternative estimate of the L skewness was obtained from a pooled frequency analysis for the catchment draining to gauging station 56013, but considering the site to be ungauged. The analysis is based on the method outlined by Kjeldsen and Jones (2009) and yields a pooled estimate of the L skewness of -0.1202, which is equivalent to a shape parameter $\kappa = 0.1202$. Similarly, the GLO scale parameter, β , was obtained from the pooled frequency analysis and estimated to be 0.202. Using the 3rd order polynomial developed in Table 1 for a record length of five years, the standard deviation of *QMED* from the short at-site sample can be determined from which the weights ω (Eq. 21) and the resulting *QMED* (39.4m³/s) and the associated *fse* (1.10) can be derived for the sparsely gauged case. In contrast, estimating the *QMED* and *fse* for the case where the standard deviation is obtained from the 3rd order polynomial for a sample size of 5, but without including the regionalised estimate of *QMED*, gives a values of *QMED* = 42.1 m³/s and *fse* = 1.22, showing that the adjustment using the added information provided by the regression model increases the precision of the estimate.

As the *fse* for the sparsely gauged case is less than the case where a regression-only estimate is adjusted using a donor site, this example highlights the value of at-site records, even if these are relatively short. In general, the results shown in Table 3 confirm the importance of including all available data into the analysis as this will help to reduce the uncertainty of the design flood estimates. In particular, data that are available at the site of interest are shown to be especially valuable.

6. DESIGN FLOOD ESTIMATES IN UNGAUGED SITES

Having focussed so far on estimation of the index flood, this section will discuss the general level of uncertainty associated with design flood estimates in ungauged sites for higher return periods (specifically 2, 5, 30 and 100 years). The analysis is based on design flood estimates obtained using the improved FEH method was undertaken based on AMAX series available from the HiFlows-UK database version 3.02 (Environment Agency, [20102009](#)). This database contains AMAX series from 955 gauging station operated by UK gauging authorities (Environment Agency, Scottish Environment Protection Agency and Rivers Agency of Northern Ireland) and located throughout the country. For each gauging station the relevant gauging authority has made an expert judgement as to the quality of the rating curve at high flows and assigned a quality rating, indicating if the data are suitable for estimation of *QMED*

and/or suitable for use in pooled analysis. For the purpose of this study, AMAX series with less than four observations were removed and only stations considered to be at least suitable for *QMED*, and to be rural rural ($URBEXT_{2000} < 0.060$), were included. This results in a reduced sample size of 715 catchments. For each gauged catchment, a set of catchment descriptors was extracted from the FEH CD-ROM v3 (CEH, 2009) enabling calculation of design flood estimates at each site as if it was ungauged. A summary of the AMAX dataset and the associated catchment descriptors are shown in Table 4 and Table 5.

TABLE 4

TABLE 5

The factorial standard error of design flood estimates for a range of return periods is described (as outlined in the appendix) using the log-transformed residuals of the T -year estimates

$$e = \ln \hat{Q}_T - \ln Q_T \quad (24)$$

Here Q_T is the design peak flow value with a T -year return period obtained by fitting a GLO distribution directly to the at-site data, and \hat{Q}_T is the corresponding estimate using the improved FEH procedure as if a site was ungauged (using either the regression model on its own or in combination with data transfer from a donor site). The factorial standard error of these ungauged estimates is estimated via the standard deviation of the residuals defined in Eq. (24) for four different return periods: 2, 5, 30 and 100 years. As the sampling error associated with each at-site design flood estimate will vary when considering extrapolation to different return periods, it was necessary to first remove this noise component before a meaningful comparison of fse values could be made across return periods. For each site the sampling noise was estimated using a Monte Carlo simulation from the estimated distribution at each site and subtracted from the squared error. Further details of this procedure are provided in the Appendix.

TABLE 6 (Results)

The factorial standard errors reported for different return periods in Table 6 show that the use of donor adjustment reduces the prediction uncertainty when compared to using the catchment descriptors-only method for estimating the index flood at ungauged sites. For example, the fse of the 100 year event when estimated using data transfer is lower than the fse of the 30 year event when estimated using catchment descriptors only. The results indicate that the 95% confidence interval of the 100 year flood obtained in an ungauged site using donor transfer will be of the order of minus 55% to plus 125% derived from using the fse values reported in Table 6 with the confidence intervals discussed in Section 3. This is the level of uncertainty that is associated with current methods, the level of data availability and the precision of the QMED equation in Eq. (12). For the 95% confidence limit on the 100 year flood estimated in an ungauged site with the support of data transfer from a donor site ($fse = 1.50$) this level of uncertainty represent a ratio of 5 between the upper and lower design estimate. If no data transfer is used, this ratio increases to more than 5.6.

It should be noted that the estimate of fse for a design event with a return period of 2 years shown in Table 6 is slightly larger than the corresponding estimate quoted in Eq. (13) in connection with the QMED equation. The estimate in Table 6 has been derived using a more updated, and slightly amended, version of the HiFlows-UK dataset than the original dataset used for the development of Eq. (13) as reported by the Environment Agency (2008). However, the difference is considered sufficiently small that the results in Table 6 can be used as a broad assessment of uncertainty levels, while a more detailed site specific investigation will have to rely on more sophisticated methods such as those provided by Kjeldsen and Jones (2006).

7. CONCLUSIONS

The aim of this study was to develop a framework for uncertainty estimation that can be applied in combination with the methods outlined in the improved FEH methodology (Environment Agency 2008) for estimating design flood events. The analytical solutions derived for a range of situations concerning data availability will enable hydrological engineers to assess the uncertainty in estimates of the index flood, thus assist in making better decisions based on available data. The framework shows that the use of local data and short

records available at a particular site should always be included in the estimation when possible.

The performance assessment at ungauged sites for a range of return periods shows that there is still considerable uncertainty associated with predictions made in ungauged catchments, but that use of local data can help to reduce the uncertainties. In general the uncertainty of the design flood estimates is large, and further research is need to identify new data and models that can help to constrain the current levels of uncertainty. In particular, the use of additional data sources beyond the annual maximum series should be explored, and the implications for uncertainty levels in design flood estimates assessed. For example, the use of peaks-over-threshold (POT) data and historical documentation of large events that might have occurred before the installation of gauging structures are both data sources routinely used in applied flood hydrology.

Acknowledgements The support provided for this research while the author was employed by the Centre for Ecology & Hydrology is acknowledged. The author is grateful to two anonymous reviewers for their helpful comments on an earlier version of the paper.

References

- Boorman, D. B., Hollis, J. M. and Lilly, A. (1995) Hydrology of Soil Types: a hydrologically-based classification of the soils of the United Kingdom. *Institute of Hydrology Report No. 126*, Institute of Hydrology, Wallingford, UK, pp. 146.
- Burn, D. H. (1990) Evaluation of regional flood frequency analysis with a region of influence approach, *Water Resour. Res.*, **26**(10), 2257–2265, doi:10.1029/WR026i010p02257.
- Castellarin, A, Kohnova, S, Gaal, L, Fleig, A, Salinas, JL, Toumazis, A, Kjeldsen TR, Macdonald, N (editors) (2012) Review of applied statistical methods for flood frequency analysis in Europe. *Milestone Report for WG2 of COST Action ES0901*, ISBN 9781906698324, 142 pages.
- Centre for Ecology & Hydrology (CEH) (2009) *FEH CD-ROM v.3*, Centre for Ecology & Hydrology, Wallingford, UK.

- Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. Chapman and Hall, London.
- Dalrymple, T. (1960) Flood frequency analysis. Water Supply Paper 1543-A, US Geol. Survey, Reston, Virginia, USA.
- Environment Agency (2008) The improved FEH statistical method. *Environment Agency R&D report SC050050/TR*, Environment Agency, Bristol, UK.
- Environment Agency (2009) HiFlows-UK database version 3.02. <http://www.environment-agency.gov.uk/hiflows/91727.aspx>
- Hall, J.W. (2011) How safe is safe enough? Editorial, *Journal of Flood Risk Management*, **4**(4), 271-272.
- Institute of Hydrology (1999) Flood Estimation Handbook (FEH), 5 Volumes, Institute of Hydrology, Wallingford, UK.
- Hosking, J. R. M., and J. R. Wallis (1993) Some statistics useful in regional frequency analysis. *Water Resour. Res.*, **29**(2), 271–281, doi:10.1029/92WR01980.
- Kjeldsen, T. R. and Jones, D. A. (2004) Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of L-moments. *Hydrology and Earth System Sciences*, **8**(2), 183-190.
- Kjeldsen, T. R. and Jones D. A. (2006) Prediction uncertainty in a median based index flood method using L-moments. *Water Resources Research*, **42**, W07414, doi :10.1029 / 2005WR004069.
- Kjeldsen T. R. and Jones, D. A. (2007) Estimation of the index flood using data transfer in the UK. *Hydrological Sciences Journal*, **52**(1), 86-98.
- Kjeldsen, T. R. and Jones, D. A. (2009) A formal statistical model for pooled analysis of extreme floods. *Hydrology Research*, **40**(5), 465-480, doi: 10.2166/nh.2009.055.
- Kjeldsen, T. R., Lamb, R. and Blazkova, S. (2013) Uncertainty in flood frequency estimation. Chapter 8 in *Applied Uncertainty Estimation for Flood Risk Management* (Eds. K. J. Beven and J. W. Hall), Imperial College Press, 500pp, London, UK.

Mkhandi, S. H., Kachroo, R. K. and Gunasekara, T. A. G. (2000) Flood frequency analysis of southern Africa: II. Identification of regional distributions. *Hydrological Sciences Journal*, **45**(3), 449-464.

Natural Environment Research Council (NERC) (1975) *Flood Studies Report*, Natural Environment Research Council, London.

Pappenberger, F. and Beven, K. J. (2006) Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, **42**, W05302, doi:10.1029/2005WR004820

Reed, D. W. (2002) Reinforcing flood-risk estimation. *Phil. Trans. R. Soc. Lond. A*, **360**(1796), 1373-1387; doi:10.1098/rsta.2002.1005

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993) *Frequency analysis of extreme events*, in Handbook of Hydrology, edited by D. Maidment, chap. 18, pp. 18.1–18.68, McGraw-Hill, New York.

Wallingford Hydro Solutions (WHS) (2009) WINFAP-FEH version 3. Wallingford Hydro Solutions, Wallingford, UK.

Yang, T., Xu, C. Y., Shao, Q. X. and Shen, X. (2010) Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach. *Stochastic Environmental Research and Risk Assessment*, **24**(2), 165-182.

Appendix: Data-based estimation of fse for ungauged sites

At each site where AMAX data are available, the T -year event has been estimated as if the site is ungauged, \hat{Q}_T , and the result compared to the T -year event derived from fitting a GLO distribution directly to the available at-site data, Q_T . The difference between the two estimates is the raw residual and it is defined as

$$e = \ln \hat{Q}_T - \ln Q_T. \quad (A1)$$

A statistical model can be formulated where the at-site estimate, Q_T , is expressed as a combination of a true value, θ , of Q_T plus an at-site sampling error, ε , which is assumed IID with zero mean, and will depend primarily on the record-length

$$\ln Q_T = \theta + \varepsilon \quad (A2)$$

A possible interpretation of θ is the value of Q_T that would be calculated if an infinite number of observations were available at the site of interest. Combining the two equations above gives

$$e = \ln \hat{Q}_T - \theta - \varepsilon \quad (A3)$$

Next, the mean-value of the squared residuals is defined from Eq. (A3) as

$$E\{e^2\} = E\{(\ln \hat{Q}_T - \theta - \varepsilon)^2\} \approx E\{(\ln \hat{Q}_T - \theta)^2\} + E\{\varepsilon^2\} \quad (A4)$$

where independence between the predicted value \hat{Q}_T and the at-site sampling error has been assumed. The expression in Eq. (A4) cannot be evaluated directly, but the corresponding mean value of squared differences that can be calculated from data is defined directly from Eq. (A1) as

$$E\{e^2\} = E\{(\ln \hat{Q}_T - \ln Q_T)^2\} \quad (A5)$$

By combining the two expressions, Eq. (A4) and (A5), for the mean value of the squared residuals, the resulting relationship is

$$E\{(\ln \hat{Q}_T - \theta)^2\} = E\{(\ln \hat{Q}_T - \ln Q_T)^2\} - E\{\varepsilon^2\} \quad (A6)$$

The variance of the at-site estimate, $E\{\varepsilon^2\}$, can be obtained from Monte Carlo simulation from the fitted GLO distribution. Assuming the errors to be identically distributed across all sites, the factorial standard error (fse) is estimated as

$$fse = \exp \left[\sqrt{\frac{1}{m} \sum_{i=1}^m (\ln \hat{Q}_T - \ln Q_T)^2 - \frac{1}{m} \sum_{i=1}^m \text{var}\{\varepsilon^2\}} \right] \quad (\text{A7})$$

Where m is the number of catchments and the variance of the at-site errors is estimated for a range of return periods by sampling 10000 AMAX series of the same length as the historical record per station, using a GLO distribution with parameters set to at site sampling values.

TABLES

Table 1: coefficients of 3rd order polynomial (Eq. 11) fitted to transformed standard deviations

Sample size	Intercept, $a_{(0)}$	1 st coeff, $a_{(1)}$	2 nd coeff, $a_{(2)}$	3 rd coeff, $a_{(3)}$
5	2.157487	1.30919296	2.08023398	-0.31107445
7	2.082885	0.7820407	1.83053293	0.94323667
10	1.965091	0.40201857	1.07072985	0.61606503
15	2.034326	0.2748751	0.53949309	0.16725382
25	1.978065	0.18330779	0.39266609	0.11129511
50	1.991291	0.09021512	0.13240444	-0.0620146
100	1.995344	0.05092219	0.05918512	-0.09961387

Table 2: catchment descriptors for subject (56013) and donor (56003) sites

Catchment descriptor	Subject site	Donor site
Easting (centroid) [m]	297622	302454
Northing (centroid) [m]	238444	237136
AREA [km ²]	63.27	62.5
SAAR [mm]	1299	1171
FARL [-]	1	0.999
BFIHOST [-]	0.494	0.528

Table 3: Estimates of QMED at subject site (location of gauging station 56013)

Method	<i>QMED</i> (m ³ /s)	<i>fse</i>
Gauged (n = 36)	36.6	1.08
Regression only	31.6	1.43
Regression + donor transfer	31.3	1.37
Sparsely gauged case (n=5 years) + regression	39.4	1.10
Sparely gauged case (n=5) without regression	42.1	1.20

Table 4: Summary of AMAX series from rural catchments in the HiFlows-UK v3.02 database

Number of gauges	715
Shortest record length	4
Longest record length	123

Mean record length	37.0
Total number of AMAX events	26487

Table 5: Summary of catchment descriptors for 715 rural catchments in the HiFlows-UK v3.02 database

Descriptor	Min.	25% percentile	50% percentile	75% percentile	Max.
AREA (km ²)	1.63	66.2	152.4	327.4	6865.0
SAAR (mm)	558	795	1016	1348	2913
FARL (-)	0.645	0.959	0.984	0.997	1.000
BFIHOST (-)	0.196	0.402	0.470	0.570	0.974
FPEXT (-)	0.002	0.035	0.052	0.078	0.295

Table 6: Factorial standard errors (*fse*) for prediction at ungauged sites with and without the use of donor transfer

Return period	fse (regression only)	fse (regression + donor)
2	1.47	1.42
5	1.48	1.43
30	1.52	1.47
100	1.54	1.50

FIGURES

Figure 1: Fitted 3rd order polynomial (Eq. 11) to standardised variance of log-median obtained from Monte Carlo simulation. Numbers next to each curve represent sample size (number of AMAX events).

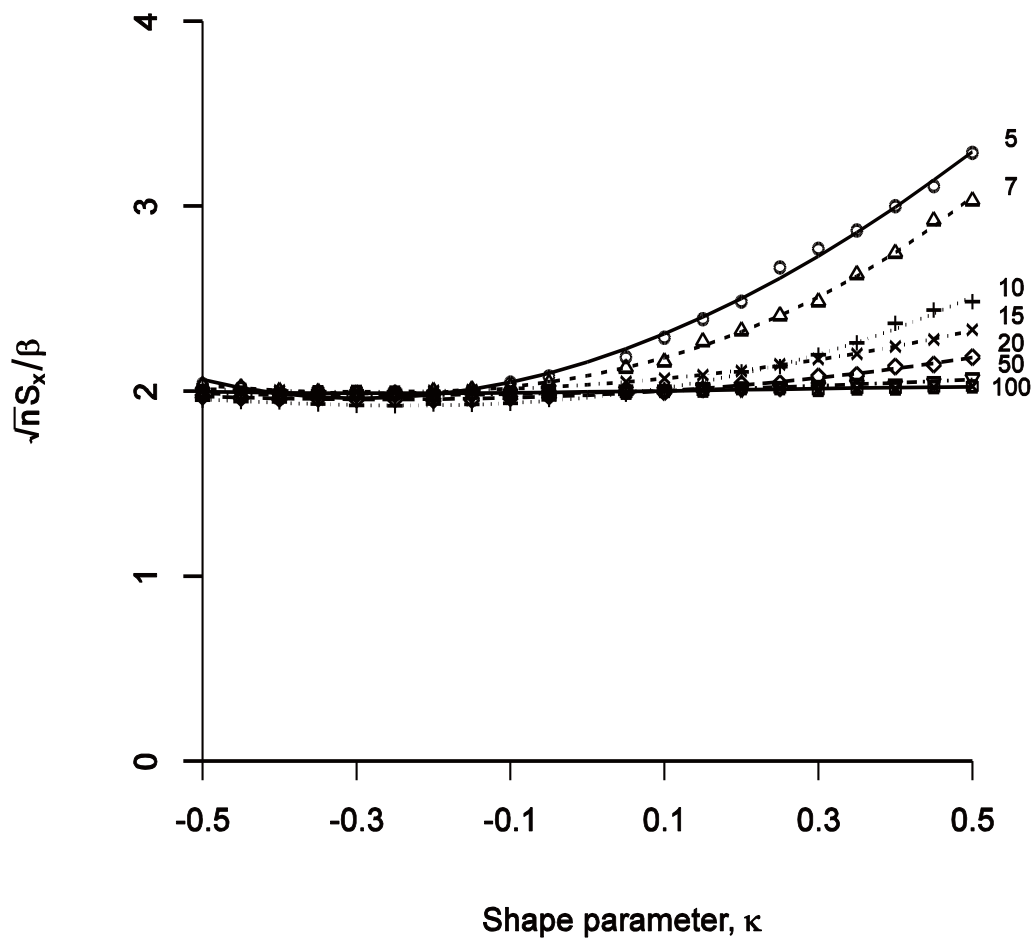


Figure 2: Annual maximum peak flow events recorded at gauging station 56013.

